Proceedings of the Second National Research Symposium on Limited English Proficient Student Issues: Focus on Evaluation and Measurement. OBEMLA, 1992.

Improving Bilingual Education Programs Through Evaluation

Alan L. Ginsburg
U.S. Department of Education

INTRODUCTION

This paper identifies how evaluation and research can contribute to federal policies for language-minority students who have limited proficiency in English. Federal bilingual education policies are controversial, and the debate benefits from systematically collected information about, and impartial evaluation of, bilingual programs.

Many of the concerns about bilingual education are misguided. For example, some people fear that bilingual policies will produce the divisive separatism that characterizes some multilingual nations, but research demonstrates little foundation for such fears in the United States. As Pease-Alvarez and Hakuta (1992) point out, by the third generation, immigrants, regardless of their nationality, have assimilated into American society, losing the ability to speak their home language fluently.

A related concern, also without foundation, is that bilingual education programs propose to substitute the home language for English. But the issue of whether immigrant children should learn English is not in question. All sides agree that English-language proficiency is essential for survival and gainful employment in America.

Concern about what instructional methods are best for achieving English-language proficiency, however, is valid. Historically, education was important to new immigrants, but not essential for success. Heavy industry and manufacturing provided plentiful well-paying jobs for people willing to do the hard work. Immigrant populations made their way upward through such employment. For America in the 1990s, however, relatively few high wage, unskilled jobs exist, and their number is continually shrinking.

The amount of education that is required to obtain a good job is increasing; a high school diploma is no longer adequate. During the 1980s real earnings declined for workers with no more than a high school education. In 1973, among men 25 to 34 years old, high school graduates earned 87 percent of what college graduates earned but, in 1987, they earned only 67 percent as much as college graduates. Some postsecondary education has now become almost as essential for well-paying jobs as a high school diploma was 20 years ago.

Tougher requirements for high school make it all the more important that language-minority students receive adequate educational opportunity. The decade of reports and piecemeal reforms since publication of A Nation At Risk has produced little gain in the educational performance of American students relative to those of other industrialized nations. Current initiatives would replace these fragmented efforts with

systemic reforms built around national education standards and national examinations. These changes would move the United States closer to the apparently more successful educational systems of our economic competitors and, if experience is any guide, would probably benefit language-minority and other students at risk for failure in school.

Education is essential to the economic success of language minorities, but the successful education of these people and other nontraditional populations is critical for our nation's economic well-being, too. Ethnic and racial minorities will account for about 30 percent of new labor force entrants over the next decade. Moreover, as the United States seeks to compete with other nations, the ability to understand and speak other languages becomes a resource to be developed.

This paper examines issues of evaluation and assessment in language-minority education within this broader context of education and its influence on the nation's future. The discussion is divided into three parts. The first part examines what has been learned from the evaluations of bilingual education conducted by the federal government during the 1980s. The second part assesses the implications of national standards and examinations for language-minorities. The final part considers how the evaluation findings and the national standard movement can suggest principles for design of future federal policies.

Evaluations During the 1980s

Background

The U.S. Supreme Court in *Lau v. Nichols* ruled that the failure to provide special language instruction to non-English speaking students (in this instance, a Chinese-speaking student) violated Title VI of the Civil Rights Act of 1964. The debate over which method of language instruction could best meet the Supreme Court requirements under *Lau* shaped the debate over bilingual education policy during the 1980s.

The Supreme Court ruling disallowed "submersion," a policy that placed children with limited English proficiency (LEP) in regular English-speaking classrooms to sink or swim, with no program to address their special educational needs. However, the Court declined to place limits on the kinds of special education services that would constitute acceptable remedies. A range of remedies might be acceptable: "Teaching English to the students of Chinese ancestry is one choice. Giving instruction to this group in Chinese is another. There may be others."

The *Lau* remedies proposed by the federal government at the close of the Carter administration sought to further clarify school district responsibilities to LEP children. Under this proposal, school systems were to assess the relative proficiency of language-minority students in English and their native language. Instruction, at least in elementary schools, would have to be provided through a student's stronger language. Although the Reagan administration withdrew the proposed regulations shortly after entering office, the deep-seated divisions over the proposed rules pointed clearly to the need for studies to evaluate systematically and rigorously the merits of alternative approaches to language instruction.

After the withdrawal of the *Lau* remedies, the national debate shifted to Title VII of the Elementary and Secondary Education Act. This legislation specifically aimed to make students proficient in the English language. But the legislation also recognized the importance of instruction in the native or dominant language "to the extent necessary to allow students to achieve competence in the English Language."

To help inform the debate, the Department of Education's Office of Planning, Budget and Evaluation

conducted a review of the literature that, far from settling the issue, fueled the controversy. The Department's report of its findings, written by Baker and de Kanter (1981) systemically assessed the quality of evaluations of bilingual education programs against a set of generally applied criteria for methodological soundness. The assessment found that few evaluations met rigorous methodological standards. The few methodologically acceptable studies seemed to show mixed results, in the sense that several different approaches could work and no approaches worked all the time. (Cziko [1992] provides a succinct survey of seven major evaluations of bilingual education.)

One of the most controversial findings in the Baker-de Kanter report was that several of the studies supported the potential effectiveness of English-language "immersion" programs. These programs taught children in English using teachers who understood the children's home language. In highlighting the immersion strategy, the Baker-de Kanter review was interpreted as advocating an all-English approach.

Longitudinal Study of Bilingual Education

The Department of Education sought to improve the quality of the evaluation of bilingual education programs by launching a multiyear plan to explore different facets of the federal role in bilingual education. The centerpiece of this plan was a rigorous longitudinal evaluation of three approaches to helping students who speak a language other than English (Ramirez, Yuen, Ramey & Pasta, 1990). The three approaches represented different degrees of exposure to English-language instruction, each reflecting a different philosophy for helping LEP students move into English-language classrooms.

In English-language immersion programs, the teacher uses English for all instruction while using the home language informally, as for occasional clarification or directions. The teacher obviously needs a working understanding of the home language but may not be fluently bilingual. Students may use the home language in responding to the teacher or talking to each other. Pupils are mainstreamed into English classrooms as soon as they have shown adequate proficiency in English.

"Late-exit" transitional programs are designed to help students become proficient in their home language before they develop proficiency in English. The teacher is fluent in both languages. Children entering elementary school receive several years of instruction in the home language. At about the fourth grade the instruction shifts gradually toward English. Students are not mainstreamed into the regular English classroom until grade 5 or 6.

The "early-exit" program is a transitional bilingual education program that is commonly used in the United States. It falls midway between the immersion and late-exit programs. Initially, instruction in the home language occurs for several hours each day, with language arts frequently taught in the native language. Content is generally taught in English. Students are mainstreamed into English-only classrooms once they have demonstrated enough mastery of English to understand the material within the regular classroom environment.

The longitudinal study by Ramirez, et al., evaluated student progress over a four-year period for students in English immersion and early exit programs and over the equivalent of six years for students in late-exit programs. (The late-exit model, which does not emphasize English-language acquisition until the later grades, required a longer period for evaluation.) To achieve maximum comparability within cost constraints, the researchers evaluated only Spanish-language programs. Although the study focused on a summative evaluation, test scores were supplemented with extensive classroom observations and parental interviews.

Significant findings include the following:

- Students in all three program models demonstrated greater-than-expected gains in achievement. Although language-minority students would normally be expected to progress more slowly than other students, all three approaches enabled the students to keep pace with their peers in regular classrooms. Nonetheless, scores of language-minority students remained considerably below the norm for other students.

The pattern of English-language progress in late-exit programs differed from the others in predictable ways. Late-exit students were initially less proficient in English. By fourth grade, about half of the students exposed to English immersion and early exit instruction were rated by their teachers as good or very good in English language comprehension, compared with 40 percent of late exit students. By sixth grade, 70 percent of the late-exit students were so rated. (Comparable sixth-grade data were not collected for immersion and early exit programs because these students typically no longer received special language instruction.) Of some importance was the fact that the rate of growth for students in late-exit programs was increasing, although there is no way to project this trend to assess whether these students would actually approach grade-level norms.

- Teachers used ineffective methods of language instruction. Regardless of the method of language instruction, students had few classroom opportunities to produce language. Teachers did most of the talking in class. When students did interact with teachers, half the time they produced no language (e.g., they were listening or gesturing); when students did speak, they typically answered with simple information recall.
- Parents of students in all three bilingual programs strongly supported English-language instruction, but their preference for Spanish-language instruction was strongly associated with whether their children's program used Spanish. More than 90 percent of the parents within each type of program wanted their children to receive extra instruction in English. With respect to the home language, only 35 percent of the parents of children in immersion programs said they favored permitting Spanish to be used in the classroom, compared with half of parents of children in early-exit programs and 86 percent of the parents of children in late-exit programs. Whether parents favored a particular instructional approach because of their preference for instruction in the home language or whether their language preference was determined by the form of their children's language instruction cannot be determined from the data.

Virtually all parents (about 90 percent or more, regardless of the type of program) want bilingual teachers. This finding may reflect the parents' preference for teachers who are able to understand their children and themselves.

- Parental involvement is facilitated by instruction in the home language. More parents of children in late-exit programs monitor their children's homework (74 percent) than do parents of children in immersion or early-exit programs (53 percent). Parents may be more comfortable with teachers or better able to help their children when instruction is given primarily in the home language.
- Students typically come from environments in which both Spanish and English are spoken: this circumstance may explain why mixed-language approaches are effective. Parents of LEP children speak to each other in Spanish 86 percent of the time and to their children in Spanish 79 percent of the time. However, their children speak to their brothers

> and sisters mostly in Spanish only about 40 percent of the time. More homes receive English language newspapers than Spanish-language papers (e.g., 45 to 37 percent), children spend 84 percent of their TV-viewing time watching English-language programs and 66 percent of their record-listening time listening to English-language records. Students also come from communities in which their neighbors are as likely to use English as Spanish.

These findings suggest that focusing evaluations on determining a single best method of language instruction for non-English-speaking children was probably the wrong approach to take to evaluation. Most special language programs in the United States represent a blend of different approaches. Indeed, the study had difficulty locating either late-exit or immersion programs, and the seven immersion programs in the study were all that could be found in the entire country. The fact that all three approaches could be effective for elementary school children indicates that the most important requirement is to learn one language well. That language does not initially have to be English, so long as transition to English occurs by the third or fourth grade.

Nonetheless, the fact that students failed to catch up to expected norms suggests that other factors, including program content, need greater consideration. Exposing language-minority and other children at risk to a more challenging curriculum is one goal of advocates for stronger national academic standards.

Bilingual Education and the Movement Toward National Standards

During the 1980s the policy debate over the appropriate method for instructing LEP students shaped the evaluation process. Little attention was given to the content of what was being taught. In the 1990s, however, evaluations of programs for language-minority students will be shaped by the outcome of the policy debate over whether this country should adopt national education standards. Proposals such as those in the Education Department's AMERICA 2000 initiative call for systemic reforms; these include setting national standards that establish what students are expected to know in core subject areas.

The final report of the National Council on Education Standards and Testing (NCEST, 1991), a congressionally created body drawing bipartisan representation from Congress, the administration, governors, teachers unions, and education experts, helped move the nation toward national standards:

> In the absence of well-defined and demanding standards, education in the United States has gravitated toward *de facto* national minimum expectations, with curricula focusing on low-level reading and arithmetic skills and on small amounts of factual material in other content areas. Most current assessment methods reinforce the emphasis on these low-level skills and processing bits of information rather than on problem solving and critical thinking. The adoption of world-class standards would force the Nation to confront today's educational performance expectations that are simply too low.

The report's conclusions are consistent with the views of most Americans. Surveys demonstrate strong public support for accountability and national tests: 7 percent favor a standardized national test, 68 percent a standardized national curriculum, and 81 percent national goals and standards.

With broad public support and evidence from other industrialized nations on the effectiveness of standards, the United States is likely to move toward some system of national standards and examinations soon. The

implications of these changes for language minority students need to be carefully explored. Concerns about the fairness of tests for language-minority and other at-risk populations could be magnified under a high-stakes national examination process.

The experience with minimum competency testing indicates that standards need not have harmful effects. When these tests were instituted during the mid-1970s, there was some concern that the requirements would hold minority students back and cause more of them to drop out of high school. But trends in student performance indicate that competency standards probably worked to the benefit of students from nontraditional backgrounds.

The National Assessment of Educational Progress represents one of the best sources of consistent information on student performance since the 1970s. In 1975, only 52 percent of Hispanic 17-year-olds read at the basic proficiency level; in 1988, 73 percent did. And the proportion who read at the adept level in 1988 (24 percent) was nearly double the proportion who read at that level in 1975 (13 percent). In addition, between the mid 1970s and 1990, Hispanics' scores on the Scholastic Aptitude Test (SAT) improved by 28 points, while white students' scores declined by 9 points. Although Hispanic drop-out rates remain unacceptably high, they appear to have declined slightly since the mid 1970s.

Despite these gains the performance of Hispanic students remains below the level for white students, and the gap worsens at higher skill levels. Because competency requirements seemed to have previously benefited at-risk students, raising requirements through new national standards and encouragement could further extend these benefits.

To work, however, national standards must be perceived as fair, must seek to challenge and motivate students to improve, and must provide students with the special resources needed to improve. Ultimately, the success of a system of national standards will depend on answers to the following questions:

- When is it appropriate to test children from non-English language backgrounds? Children exposed to English for the first time presumably need a transition period before testing. Conversely, students must not be excluded from testing for so long that schools are no longer held accountable for their performance.
- In what language is it appropriate or even feasible to administer the test? Issues of feasibility, accuracy, and appropriateness have to be resolved. How feasible is it to translate tests into languages other than English, and what is the cost of doing so? Can a student's stronger language be accurately determined? Is it appropriate to test knowledge of the English language while testing knowledge of the content of other subjects in a student's stronger language?
- How can test results be used to expand student opportunities rather than simply to punish students who are experiencing difficulties? Testing can reinforce students' educational opportunities, if schools use test results to identify and correct student weaknesses. The NCEST has proposed testing students in grades 4 and 8 as well as grade 12 in order to detect and correct deficiencies. Such a test pattern would differ from the practice of most other industrialized nations, which test students only once before tracking them into college programs.

Early identification of problems, of course, does not guarantee that needy students will receive special support. Schools should be required to address special problems as a condition of testing. Moreover, if schools are failing, they should be held accountable. Many schools, particularly ones serving lower income

areas, are insulated from pressures to provide high-quality education to all children. Recent legislation included in Chapter 1 of the Elementary and Secondary Education Act requires schools that fail to meet performance goals to institute a performance improvement plan. A Chapter 1 type of improvement plan could be extended to cover schools failing language-minority students.

- How can the tested material be coordinated with a challenging curriculum? A valid criticism of current standardized testing is that the material on which students are tested may never be taught in school. This circumstance puts at-risk students at a particular disadvantage, because these students are least likely to be exposed to the range of general-knowledge questions on standardized tests. Aligning course content and tests with curriculum frameworks would give at-risk students a fairer chance.

New standards would have implications for federal evaluation requirements under Title VII. The current Title VII legislation requires local programs to report an almost impossible amount of information: subject areas taught; instructional methods; time spent on specific tasks; preparation, language abilities, and educational background of the staff; students' achievements in English language arts and subject areas, oral proficiency in English, and achievement in native language; each school's grade retention rate, dropout rate, absenteeism, number of referrals to special education, number of placements in gifted and talented programs, and postsecondary education attendance.

Faced with excessive reporting burdens, recipients of federal bilingual education grants have simply ignored most of them. A 1990 evaluation independently assessed the quality of Title VII evaluation reports. Although most programs used appropriate achievement tests, fewer programs analyzed test data appropriately. Only about half used a 12-month testing interval, although use of shorter test intervals is known to seriously overstate gains in student achievement. Less than a quarter of the programs reported test data in sufficient detail to draw programmatic conclusions. Finally, very few programs (about 15 percent) followed former participants to assess their progress in the regular education program, although this assessment may represent the best measure of program effectiveness.

Instead of being a paper exercise, local evaluations of federal bilingual education programs should become an integral part of program operations. Evaluations should focus on the performance of students in relation to national standards, and the quality of local program evaluations must improve considerably.

Implications for the Federal Role

As already mentioned, two sets of issues have been explored in the evaluation of programs for language minorities: in the 1980s, the focus was on instructional processes, while in the 1990s the focus is on instructional content. These two evaluation streams need to be combined in a coherent strategy that integrates the content of what is taught and the methods of instruction.

The upcoming re-authorization of Title VII offers an opportunity to debate and craft legislative responses that build on evaluation evidence and new educational reforms. Although the details of reform will require careful analysis, here are five general principles that could help guide reforms:

1. Bilingual programs should be held accountable for high achievement by their students, while local programs should be allowed flexibility over the method of bilingual education.
Evaluations have demonstrated that bilingual education can work, but that no one method is uniformly superior. Successful programs may focus on dual language development or may

immerse children in English immediately. In return for strong accountability for student performance, the federal government should expand local program discretion over federal resources. For instance, federal legislation discourages programs from serving students for more than three years. If student performance is satisfactory, there is no reason to limit the length of bilingual education programs.

2. Teachers of LEP students in bilingual and regular classrooms need sound training. Evaluations have shown that even teachers in thoughtfully designed programs appear to use pedagogies that are not effective. The federal government's Title VII program should become a major source of teacher training support, but this support should ensure that the training provided is sound and likely to take hold in a school. Bilingual education training, now focused almost entirely on teachers in the bilingual program, might be extended school-wide. Because all teachers in the school work with language-minority children, all could benefit from training in language instructional approaches.

3. Accountability requirements should shift from traditional standardized tests to performance-based examinations that promote opportunities for language minority and other at-risk populations to achieve "World Class" standards. Current standardized tests are not well coordinated with the curricula or services. Teachers perceive these tests as having little value and as being primarily punitive. A system of national standards tied to examinations must be linked to curricula. Poorly performing students should receive special help to enable them to reach the standards. Furthermore, schools that consistently fail such students need to be held accountable for this failure and not permitted to continue to operate on a business-as-usual basis.

Language-minority children should be excluded from testing only if they enter school with limited English proficiency, and then only for a specific period. Widespread exclusion would serve to stigmatize excluded students and diminish schools' accountability to provide the students with appropriate educational services.

4. The federal Government should launch a multiyear agenda to identify best practices within different instructional approaches, rather than attempting to determine a single best approach. The evaluations of bilingual education in the 1980s sought a single winner to the question of identifying effective methods of language instruction. This approach was wrong. Evaluations for the 1990s need to be driven by the question of what approach works best under what conditions.

Research should also focus on strategies to encourage students to learn English outside school and to foster parental involvement in their children's education. These efforts should build on evaluation findings that show that language-minority parents will become more involved in education when schools communicate with them in their home language.

5. Federal bilingual education policy should recognize that the home language is a resource to be developed. Achieving bilingualism through foreign-language instruction for native-born Americans is an accepted national priority, one that is becoming more important in an increasingly competitive economic environment. Logically it follows that students who want to maintain their home language should have the opportunity to do so. A knowledge of the home language is not a substitute for strong knowledge of English but a recognition that knowledge of the home language and of English can help the development of both languages.

Next Page

** References **
Next Research Paper
Table of Contents
Top of Page